Homework 5
CS157b
Craig Huff
009841390

20.5.2:

a) (0,1,*P),* (0,2,*P),* (2,0,*R),* (1,0,*R),* (0,2,C), (0,1, *C)*
Give an example of a sequence of messages that could occur if site 1 wants to commit and site 2 wants to abort.

(0,1,P), (0,2,P), (1,0,R), (2,0,D), (0,1,A), (0,2,A)

20.6.2:

Suppose there are five sites, each with a copy of a database element X . One of these sites P is the dominant site for X and will be used as I's primary site in a primary-copy distributed-lock system. The statistics regarding accesses to X are:

i. 50% of all accesses are read-only accesses originating at P.
ii. Each of the other four sites originates 10% of the accesses, and these are read-only.
iii. The remaining 10% of accesses require exclusive access and may originate at any of the five sites with equal probability (i.e., 2% originate at each).

For each of the lock methods below, give the average number of messages needed to obtain a lock. Assume that all requests are granted, so no denial messages are needed.

a) Read-locks-one; write-locks-all.
90% of the accesses are read only require no messages, because there is a lock table and a copy at each site. The other 10% of accesses require exclusive locks which produces three messages between the originating site and the other four sites (12 messages per lock). Thus, the average number of messages is 1.2.

b) Majority locking.
For a local lock it's possible to broadcast a message that is received by all the other sites. Locks being released can be broadcast the message. For majority locking we need at least three requests to issue the lock. So three messages to request the lock, one broadcasted message accepting the either shared lock or exclusive lock, and one broadcasted message saying the lock has been released. Thus, the average number of messages is 5.

c) Primary-copy locking, with the primary copy at P.
52% of the accesses are requesting locks are coming from the primary-copy P, and so there will be no messages generated. The remaining 48% will produce three messages between the originating site and the other four sites (12 messages per lock). Thus, the average number of messages is 5.76.

22.1.1:
Suppose we are given the eight “market baskets” of Fig. 22.3.

B_1 = {milk, coke, beer}
B_2 = {milk, pepsi, juice}
B_3= {milk, beer}
B_4= {coke, juice}
B_5= {milk, pepsi, beer}
B_6 = {milk, beer, juice, pepsi}
B_7= {coke, beer, juice}
B_8= {beer, pepsi}

Figure 22.3

a) As a percentage of the baskets, what is the support of the set {beer,juice}?
Since only B7 and B6 have both beer and juice the support for the set is 25%.

b) What is the support of the item set {coke, pepsi}?
None of the sets have coke and pepsi, resulting in 0% support.

c) What is the confidence of milk given beer (i.e., of the association rule {beer} => milk)?
Beer appears in six of the sets. Out of those six sets, milk appears in four of the sets, giving the confidence of milk given beer 67%.

d) What is the confidence of juice given milk?
Milk appears in five of the sets. Out of those five sets, juice appears in only two. Resulting in confidence of 40% of juice given milk.

e) What is the confidence of coke, given beer and juice?
{Beer, Juice} appears in two sets, with coke only appearing in one set. Resulting in confidence of 50% of coke, given beer and juice.

f) If the support threshold is 37.5% (i.e., 3 out of the eight baskets are needed), which pairs of items are frequent?
{milk, pepsi}, {milk, beer}, {pepsi, beer}

g) If the support threshold is 50%, which pairs of items are frequent?
{beer, milk} has 50% support (Sets B_1, B_3, B_5, B_6)

h) What is the most interesting association rule with a singleton set on the left?
{Coke} -> Pepsi has the highest negative interest because no purchases of coke and Pepsi in the same set.

22.2.1:
Simulate the A-Priori Algorithm on the data of Fig. 22.3, with s = 3.

	Milk	**Beer**	**Pepsi**	**Juice**
Coke	B_1	B_7	none	B_4
Juice	B_2, B_6	B_6	B_2, B_6	
Pepsi	B_2, B_5, B_6	B_5, B_6, B_8		
Beer	B_1, B_3, B_5, B_6			

Singletons: {Milk}, {Beer}, {Pepsi}, {Juice}, {Coke}
Doubletons: {Milk, Beer}, {Milk, Pepsi}, {Pepsi, Beer}
Triplets: None